



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 13, 2024 – 11:41 AM EDT

PDB ID : 1O7A  
Title : Human beta-Hexosaminidase B  
Authors : Maier, T.; Strater, N.; Schuette, C.; Klingenstein, R.; Sandhoff, K.; Saenger, W.  
Deposited on : 2002-10-29  
Resolution : 2.25 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

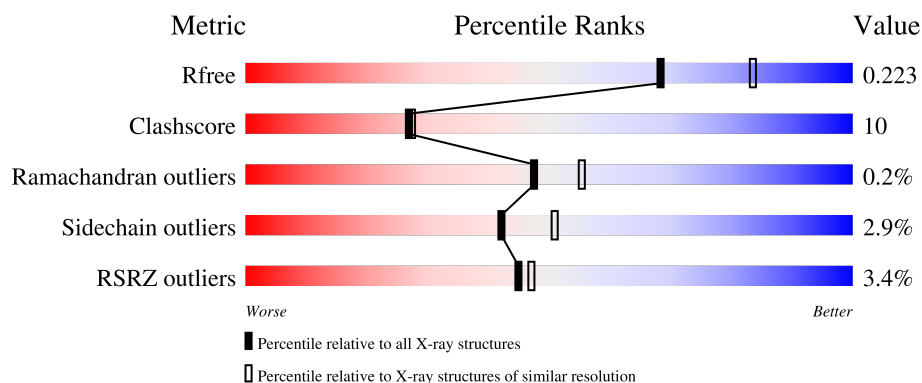
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.25 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1377 (2.26-2.26)
Clashscore	141614	1487 (2.26-2.26)
Ramachandran outliers	138981	1449 (2.26-2.26)
Sidechain outliers	138945	1450 (2.26-2.26)
RSRZ outliers	127900	1356 (2.26-2.26)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	515	<div> <div>3%</div> <div> <div></div> <div>72%</div> <div>20%</div> <div>6%</div> </div> </div>
1	B	515	<div> <div>2%</div> <div> <div></div> <div>77%</div> <div>16%</div> <div>6%</div> </div> </div>
1	C	515	<div> <div>4%</div> <div> <div></div> <div>77%</div> <div>16%</div> <div>6%</div> </div> </div>
1	D	515	<div> <div>3%</div> <div> <div></div> <div>74%</div> <div>18%</div> <div>6%</div> </div> </div>
1	E	515	<div> <div>3%</div> <div> <div></div> <div>74%</div> <div>18%</div> <div>6%</div> </div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	515	
2	G	2	
2	H	2	
2	I	2	
2	J	2	
2	K	2	
2	L	2	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	NAG	G	2	-	-	-	X
2	NAG	H	2	-	-	-	X
2	NAG	I	2	-	-	-	X
2	NAG	J	2	-	-	-	X
2	NAG	K	2	-	-	-	X
4	NAG	C	703	-	-	-	X
4	NAG	E	702	-	-	-	X
4	NAG	F	702	-	-	-	X
4	NAG	F	703	-	-	-	X
5	EDO	B	801	-	-	X	-

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 26351 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called BETA-HEXOSAMINIDASE BETA CHAIN.

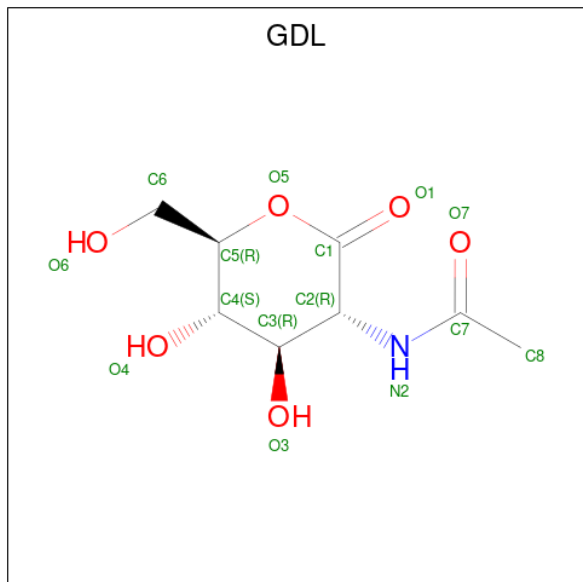
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	484	Total	C	N	O	S	0	4	1
			3933	2538	654	725	16			
1	B	484	Total	C	N	O	S	0	3	1
			3927	2535	653	724	15			
1	C	484	Total	C	N	O	S	0	4	1
			3935	2538	656	727	14			
1	D	483	Total	C	N	O	S	0	6	1
			3939	2542	655	726	16			
1	E	484	Total	C	N	O	S	0	5	1
			3941	2542	655	728	16			
1	F	484	Total	C	N	O	S	0	6	1
			3948	2546	656	730	16			

- Molecule 2 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	G	2	Total	C	N	O	0	0	0
			28	16	2	10			
2	H	2	Total	C	N	O	0	0	0
			28	16	2	10			
2	I	2	Total	C	N	O	0	0	0
			28	16	2	10			
2	J	2	Total	C	N	O	0	0	0
			28	16	2	10			
2	K	2	Total	C	N	O	0	0	0
			28	16	2	10			
2	L	2	Total	C	N	O	0	0	0
			28	16	2	10			

- Molecule 3 is 2-(acetylamido)-2-deoxy-D-glucono-1,5-lactone (three-letter code: GDL) (formula:  $C_8H_{13}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			15	8	1	6		
3	B	1	Total	C	N	O	0	0
			15	8	1	6		
3	C	1	Total	C	N	O	0	0
			15	8	1	6		
3	D	1	Total	C	N	O	0	0
			15	8	1	6		
3	E	1	Total	C	N	O	0	0
			15	8	1	6		
3	F	1	Total	C	N	O	0	0
			15	8	1	6		

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	0	0
			14	8	1	5		
4	B	1	Total	C	N	O	0	0
			14	8	1	5		
4	C	1	Total	C	N	O	0	0
			14	8	1	5		
4	E	1	Total	C	N	O	0	0
			14	8	1	5		
4	F	1	Total	C	N	O	0	0
			14	8	1	5		
4	F	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		
5	C	1	Total	C	O	0	0
			4	2	2		
5	C	1	Total	C	O	0	0
			4	2	2		
5	C	1	Total	C	O	0	0
			4	2	2		
5	E	1	Total	C	O	0	0
			4	2	2		
5	E	1	Total	C	O	0	0
			4	2	2		
5	E	1	Total	C	O	0	0
			4	2	2		
5	F	1	Total	C	O	0	0
			4	2	2		
5	F	1	Total	C	O	0	0
			4	2	2		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	F	1	Total	C	O	0	0
			4	2	2		

- Molecule 6 is water.

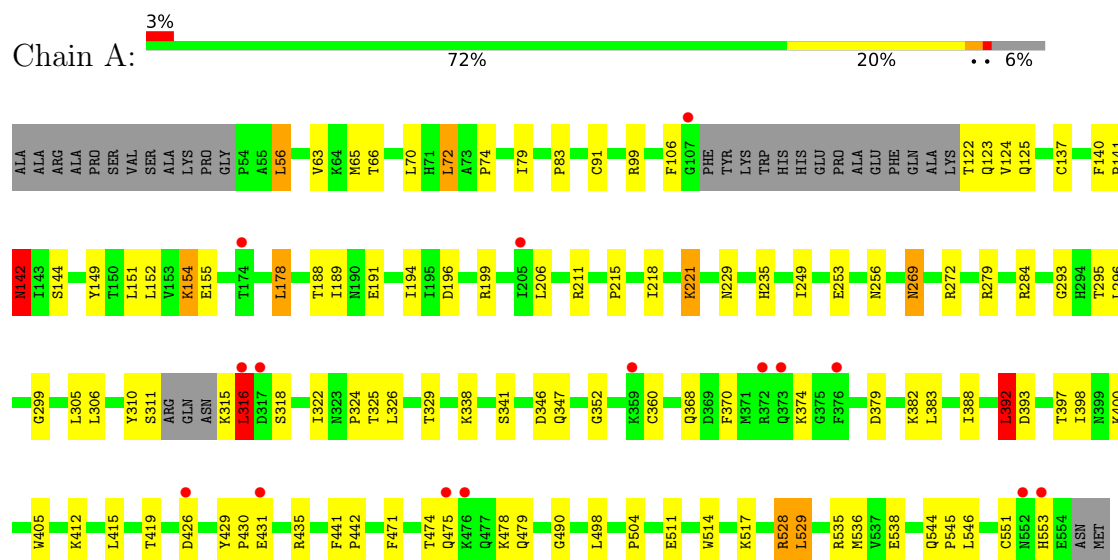
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	385	Total	O	0	0
			385	385		
6	B	399	Total	O	0	0
			399	399		
6	C	385	Total	O	0	0
			385	385		
6	D	346	Total	O	0	0
			346	346		
6	E	448	Total	O	0	0
			448	448		
6	F	363	Total	O	0	0
			363	363		



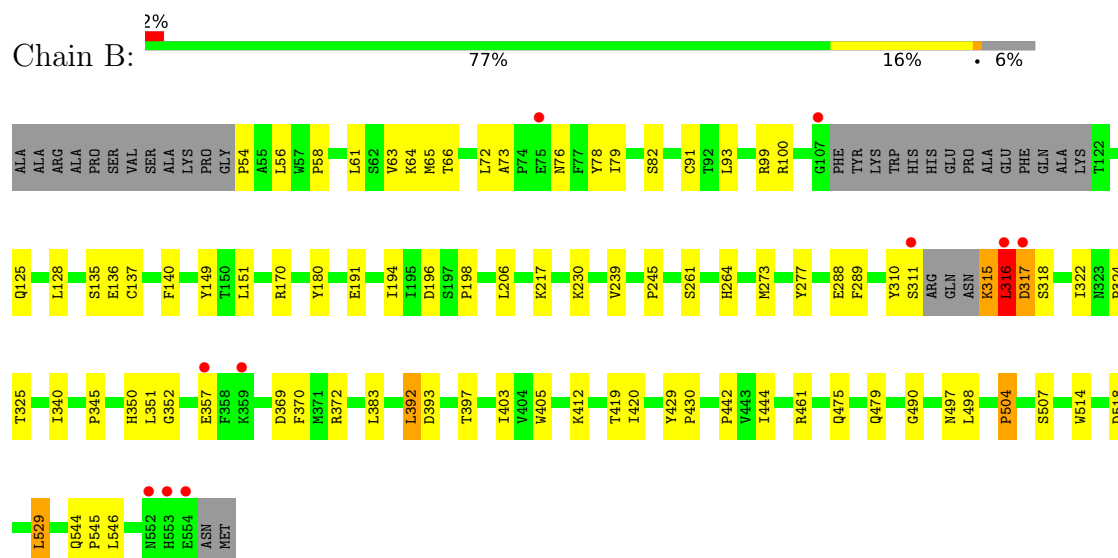
### 3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

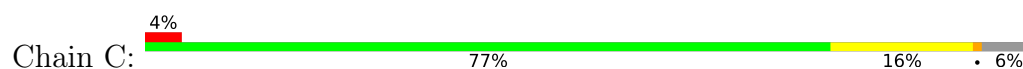
#### • Molecule 1: BETA-HEXOSAMINIDASE BETA CHAIN

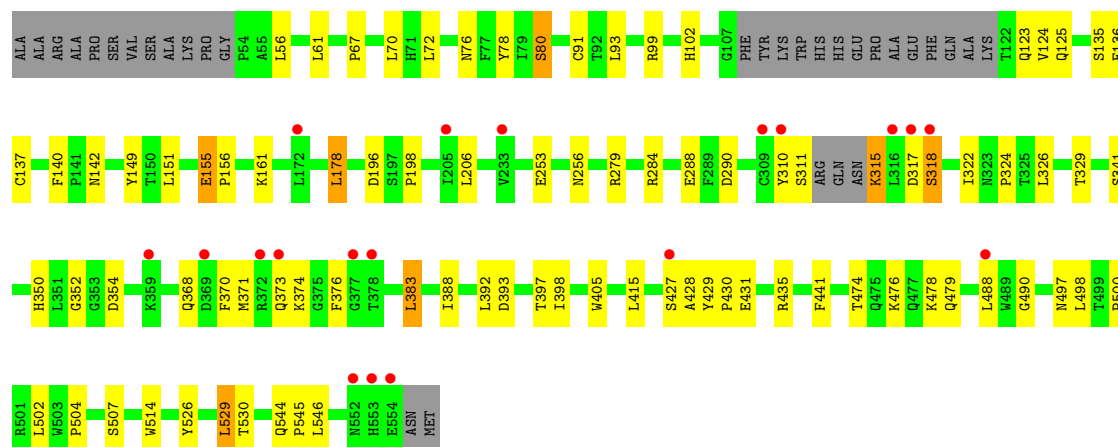


#### • Molecule 1: BETA-HEXOSAMINIDASE BETA CHAIN

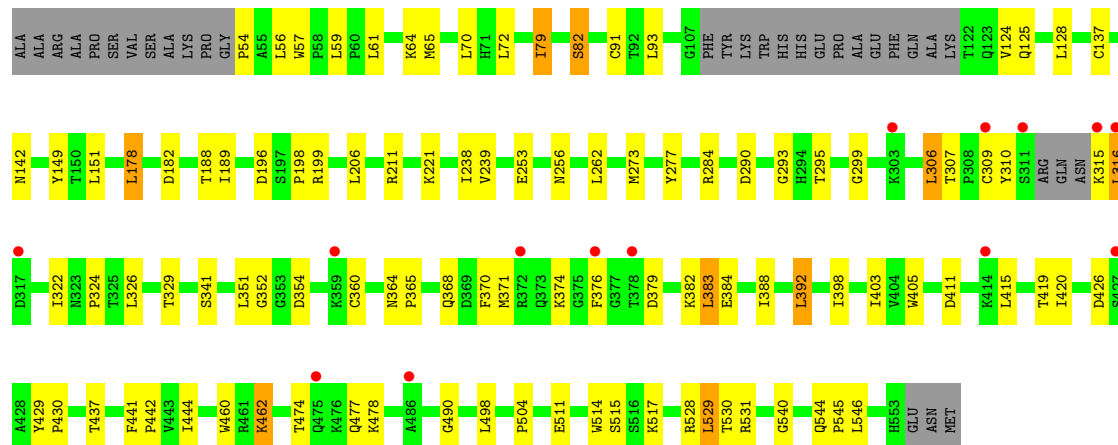
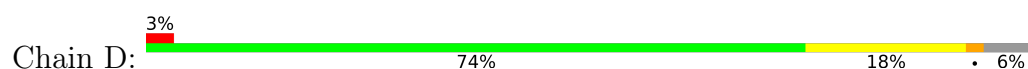


#### • Molecule 1: BETA-HEXOSAMINIDASE BETA CHAIN

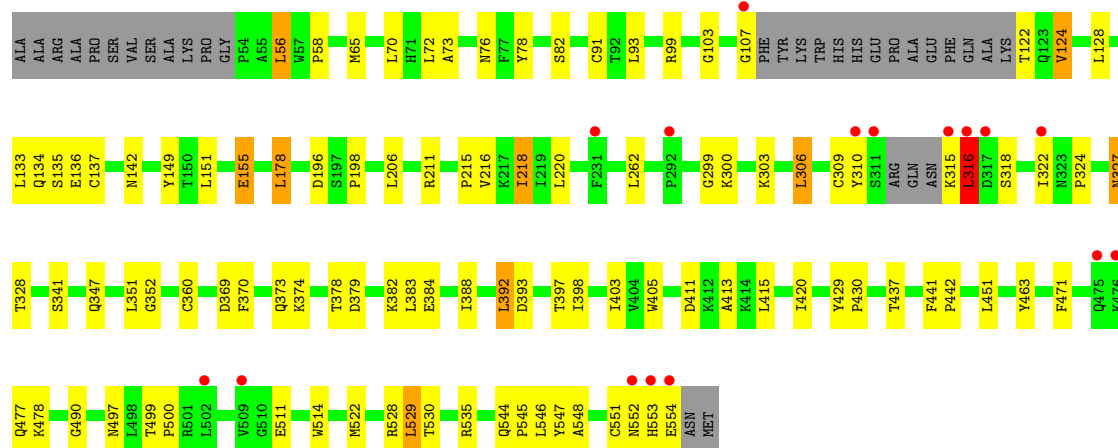
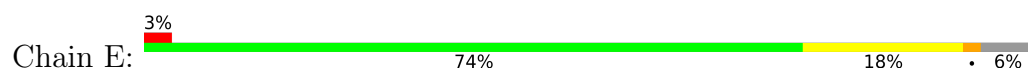




• Molecule 1: BETA-HEXOSAMINIDASE BETA CHAIN



• Molecule 1: BETA-HEXOSAMINIDASE BETA CHAIN






- Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain K:  50% 50%

MAG1  
MAG2

- Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain L:  50% 50%

MAG1  
MAG2

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	163.93Å 163.93Å 244.72Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	28.12 – 2.25 28.12 – 2.25	Depositor EDS
% Data completeness (in resolution range)	97.4 (28.12-2.25) 97.4 (28.12-2.25)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.84 (at 2.24Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, $R_{free}$	0.196 , 0.236 0.187 , 0.223	Depositor DCC
$R_{free}$ test set	1715 reflections (0.98%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	35.3	Xtriage
Anisotropy	0.622	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 47.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.53$ , $\langle L^2 \rangle = 0.37$	Xtriage
Estimated twinning fraction	0.004 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	26351	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.00% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NAG, GDL, EDO

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
1	A	0.54	1/4045 (0.0%)	0.73	6/5499 (0.1%)
1	B	0.50	0/4039	0.71	1/5491 (0.0%)
1	C	0.49	0/4047	0.68	2/5503 (0.0%)
1	D	0.46	0/4050	0.69	2/5506 (0.0%)
1	E	0.51	0/4053	0.72	3/5510 (0.1%)
1	F	0.47	0/4060	0.70	3/5520 (0.1%)
All	All	0.50	1/24294 (0.0%)	0.70	17/33029 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	142	ASN	CG-ND2	12.73	1.64	1.32

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	142	ASN	OD1-CG-ND2	-11.23	96.07	121.90
1	E	155	GLU	C-N-CD	-7.06	105.07	120.60
1	A	142	ASN	CB-CG-ND2	6.98	133.45	116.70
1	F	316	LEU	CA-CB-CG	-6.56	100.20	115.30
1	E	490	GLY	N-CA-C	5.88	127.81	113.10

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	142	ASN	Sidechain

## 5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3933	0	3832	91	0
1	B	3927	0	3828	70	0
1	C	3935	0	3830	72	0
1	D	3939	0	3843	86	0
1	E	3941	0	3835	79	0
1	F	3948	0	3840	75	0
2	G	28	0	25	2	0
2	H	28	0	25	4	0
2	I	28	0	25	4	0
2	J	28	0	25	2	0
2	K	28	0	25	2	0
2	L	28	0	25	2	0
3	A	15	0	8	0	0
3	B	15	0	8	0	0
3	C	15	0	8	0	0
3	D	15	0	8	0	0
3	E	15	0	8	0	0
3	F	15	0	8	0	0
4	A	14	0	13	1	0
4	B	14	0	13	0	0
4	C	14	0	13	0	0
4	E	14	0	13	0	0
4	F	28	0	26	0	0
5	A	12	0	18	4	0
5	B	12	0	18	4	0
5	C	12	0	18	2	0
5	E	12	0	18	1	0
5	F	12	0	18	1	0
6	A	385	0	0	17	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	B	399	0	0	11	0
6	C	385	0	0	11	0
6	D	346	0	0	8	0
6	E	448	0	0	14	0
6	F	363	0	0	11	0
All	All	26351	0	23374	479	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 10.

The worst 5 of 479 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:56:LEU:HD21	1:D:178:LEU:HD13	1.51	0.93
1:C:488:LEU:HD23	1:C:502:LEU:HD13	1.51	0.92
1:C:393:ASP:O	1:C:397:THR:HG23	1.70	0.92
1:C:488:LEU:HD12	1:C:488:LEU:O	1.74	0.87
1:F:553:HIS:ND1	1:F:554:GLU:N	2.22	0.86

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	482/515 (94%)	466 (97%)	16 (3%)	0	100	100
1	B	481/515 (93%)	462 (96%)	18 (4%)	1 (0%)	47	55
1	C	482/515 (94%)	467 (97%)	14 (3%)	1 (0%)	47	55
1	D	483/515 (94%)	462 (96%)	20 (4%)	1 (0%)	47	55
1	E	483/515 (94%)	463 (96%)	18 (4%)	2 (0%)	34	37

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	F	484/515 (94%)	463 (96%)	19 (4%)	2 (0%)	34	37
All	All	2895/3090 (94%)	2783 (96%)	105 (4%)	7 (0%)	47	55

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	316	LEU
1	C	318	SER
1	D	316	LEU
1	E	552	ASN
1	F	316	LEU

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	433/454 (95%)	418 (96%)	15 (4%)	36	43
1	B	432/454 (95%)	422 (98%)	10 (2%)	50	59
1	C	433/454 (95%)	420 (97%)	13 (3%)	41	50
1	D	434/454 (96%)	422 (97%)	12 (3%)	43	52
1	E	434/454 (96%)	421 (97%)	13 (3%)	41	50
1	F	435/454 (96%)	423 (97%)	12 (3%)	43	52
All	All	2601/2724 (96%)	2526 (97%)	75 (3%)	42	51

5 of 75 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	316	LEU
1	F	383	LEU
1	E	383	LEU
1	F	151	LEU
1	B	529	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	479	GLN
1	E	327	ASN
1	F	142	ASN
1	F	102	HIS
1	E	76	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

12 monosaccharides are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
2	NAG	G	1	1,2	14,14,15	0.51	0	17,19,21	0.92	0
2	NAG	G	2	2	14,14,15	0.54	0	17,19,21	0.71	0
2	NAG	H	1	1,2	14,14,15	0.42	0	17,19,21	0.68	0
2	NAG	H	2	2	14,14,15	0.52	0	17,19,21	0.66	0
2	NAG	I	1	1,2	14,14,15	0.43	0	17,19,21	0.88	1 (5%)
2	NAG	I	2	2	14,14,15	0.56	0	17,19,21	0.67	0
2	NAG	J	1	1,2	14,14,15	0.49	0	17,19,21	0.92	1 (5%)
2	NAG	J	2	2	14,14,15	0.54	0	17,19,21	0.78	1 (5%)
2	NAG	K	1	1,2	14,14,15	0.47	0	17,19,21	1.11	2 (11%)
2	NAG	K	2	2	14,14,15	0.52	0	17,19,21	0.67	0
2	NAG	L	1	1,2	14,14,15	0.52	0	17,19,21	0.80	1 (5%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	NAG	L	2	2	14,14,15	0.63	0	17,19,21	0.62	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NAG	G	1	1,2	-	2/6/23/26	0/1/1/1
2	NAG	G	2	2	-	6/6/23/26	0/1/1/1
2	NAG	H	1	1,2	-	2/6/23/26	0/1/1/1
2	NAG	H	2	2	-	4/6/23/26	0/1/1/1
2	NAG	I	1	1,2	-	4/6/23/26	0/1/1/1
2	NAG	I	2	2	-	4/6/23/26	0/1/1/1
2	NAG	J	1	1,2	-	3/6/23/26	0/1/1/1
2	NAG	J	2	2	-	4/6/23/26	0/1/1/1
2	NAG	K	1	1,2	-	2/6/23/26	0/1/1/1
2	NAG	K	2	2	-	5/6/23/26	0/1/1/1
2	NAG	L	1	1,2	-	4/6/23/26	0/1/1/1
2	NAG	L	2	2	-	4/6/23/26	0/1/1/1

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	K	1	NAG	C3-C4-C5	2.60	114.88	110.24
2	J	1	NAG	C2-N2-C7	-2.58	119.23	122.90
2	L	1	NAG	C2-N2-C7	-2.23	119.72	122.90
2	J	2	NAG	C2-N2-C7	-2.20	119.78	122.90
2	K	1	NAG	C2-N2-C7	-2.10	119.91	122.90

There are no chirality outliers.

5 of 44 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	G	2	NAG	C8-C7-N2-C2
2	G	2	NAG	O7-C7-N2-C2
2	H	2	NAG	C8-C7-N2-C2

*Continued on next page...*

*Continued from previous page...*

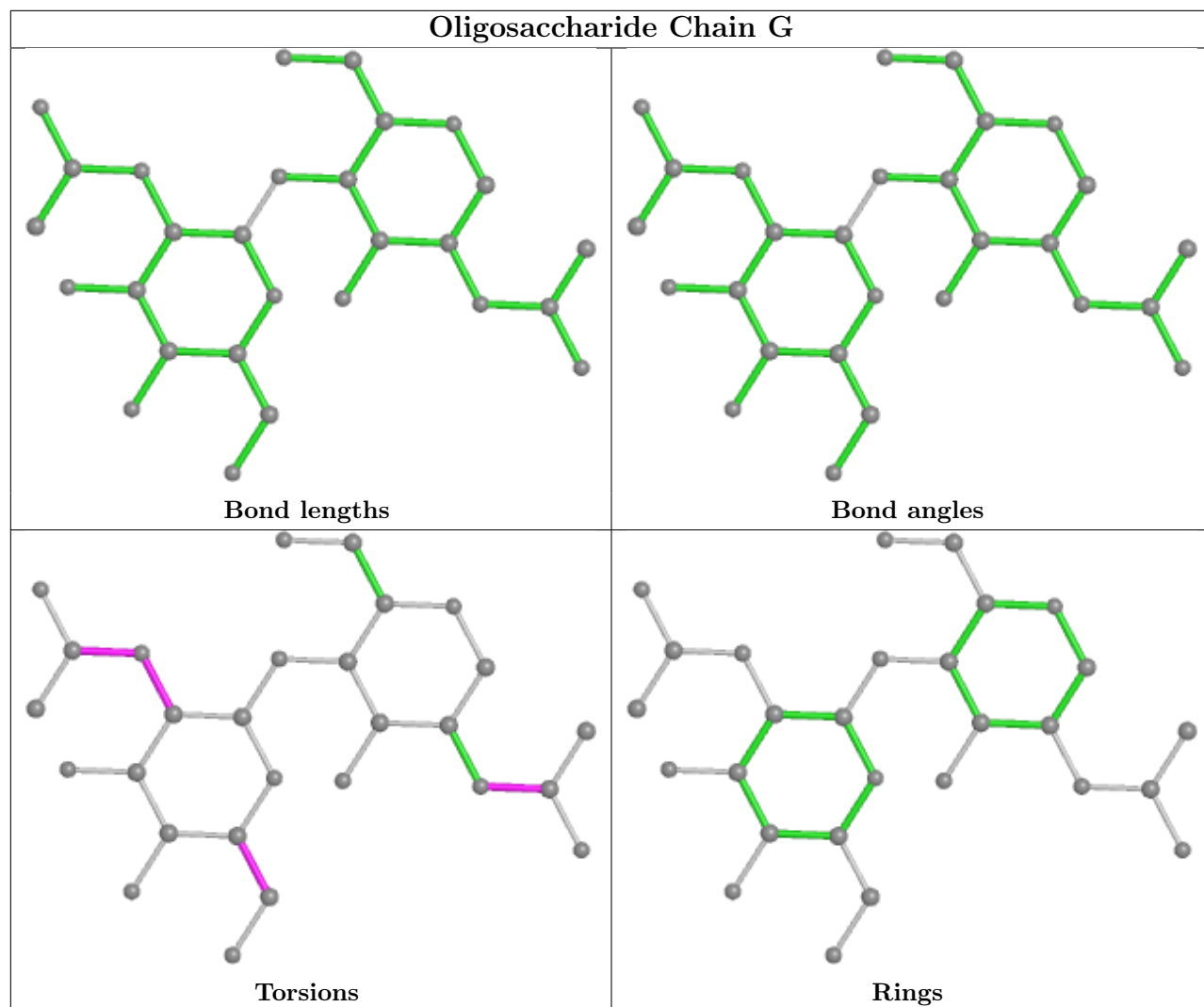
Mol	Chain	Res	Type	Atoms
2	H	2	NAG	O7-C7-N2-C2
2	I	2	NAG	C8-C7-N2-C2

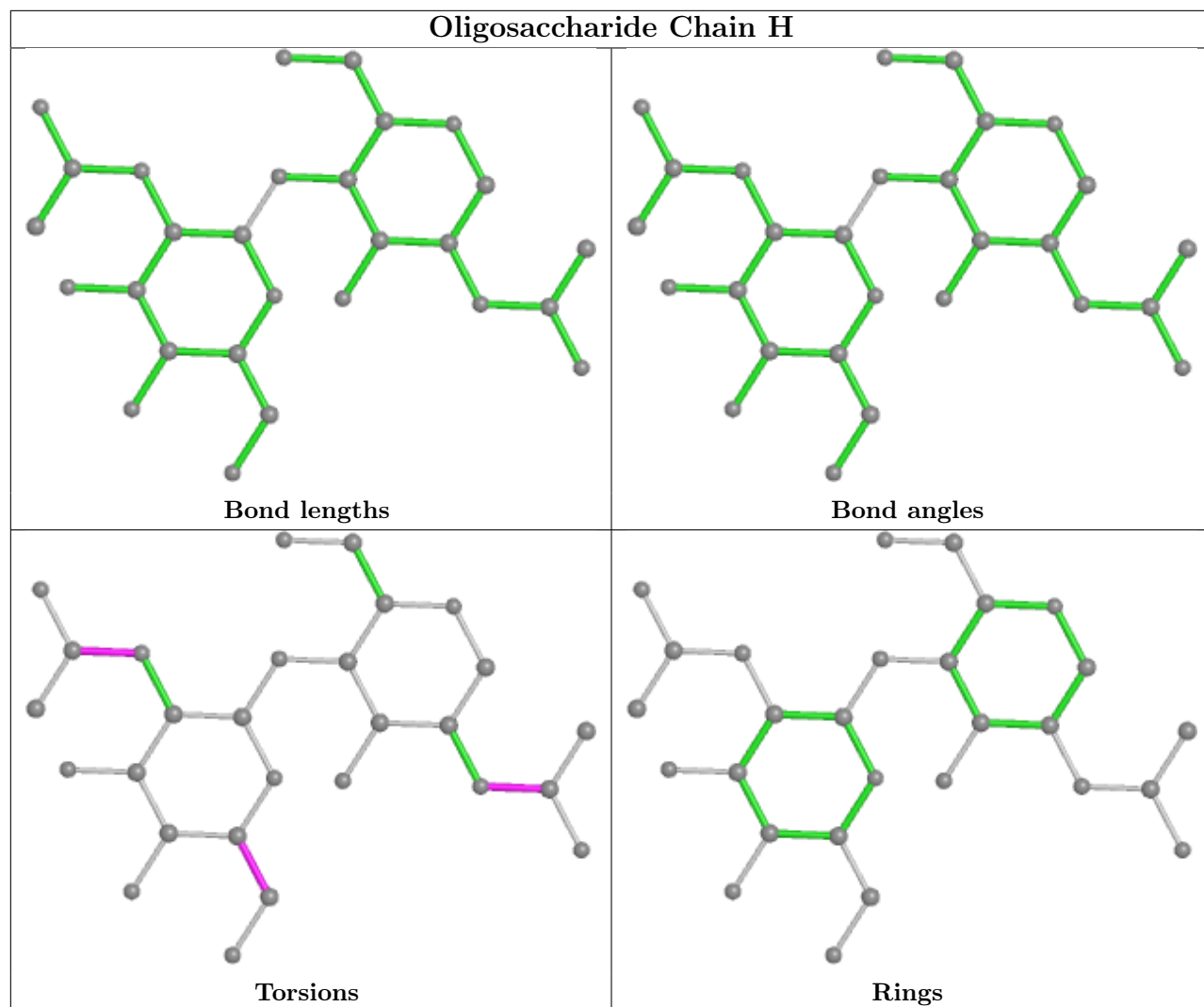
There are no ring outliers.

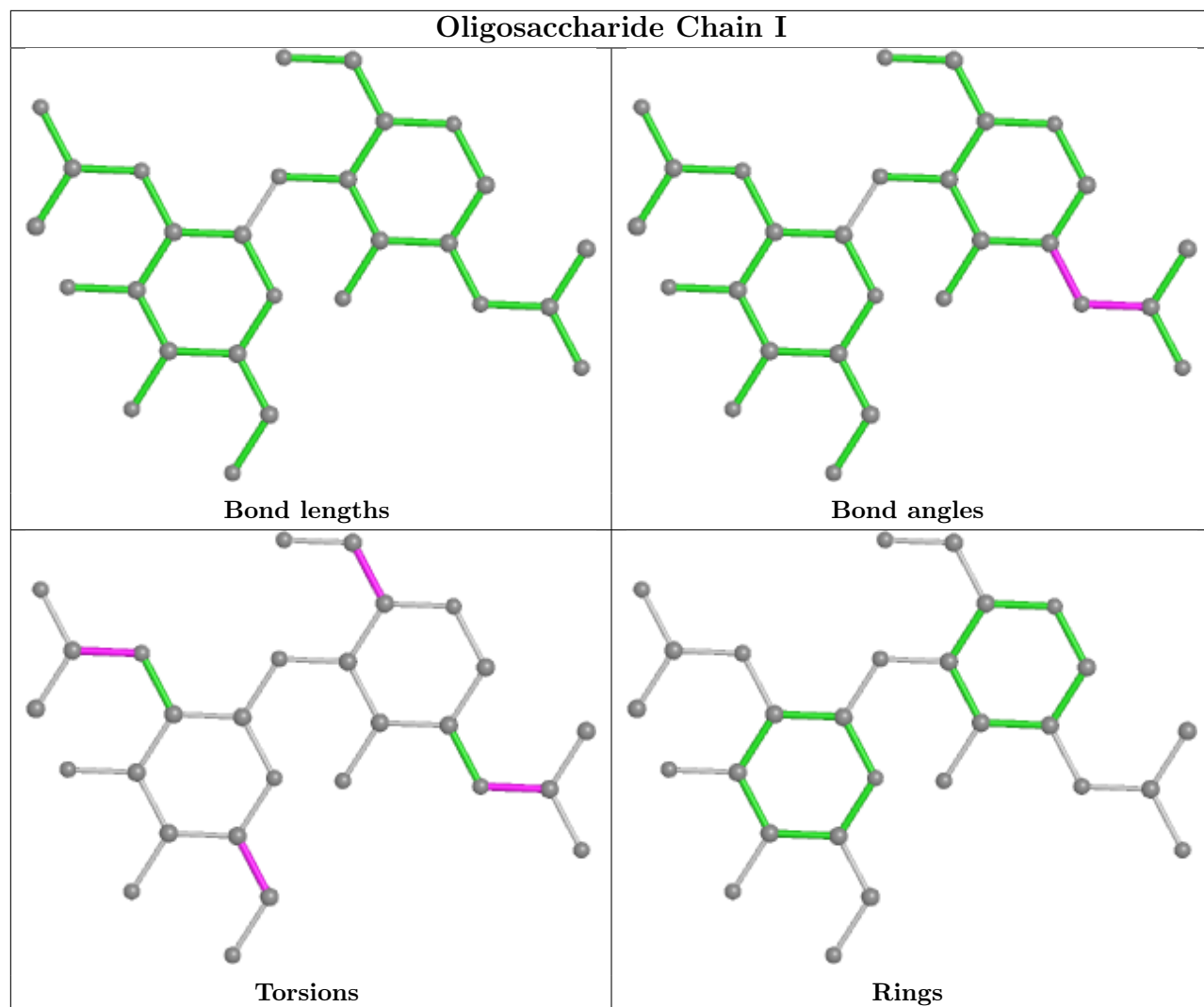
11 monomers are involved in 16 short contacts:

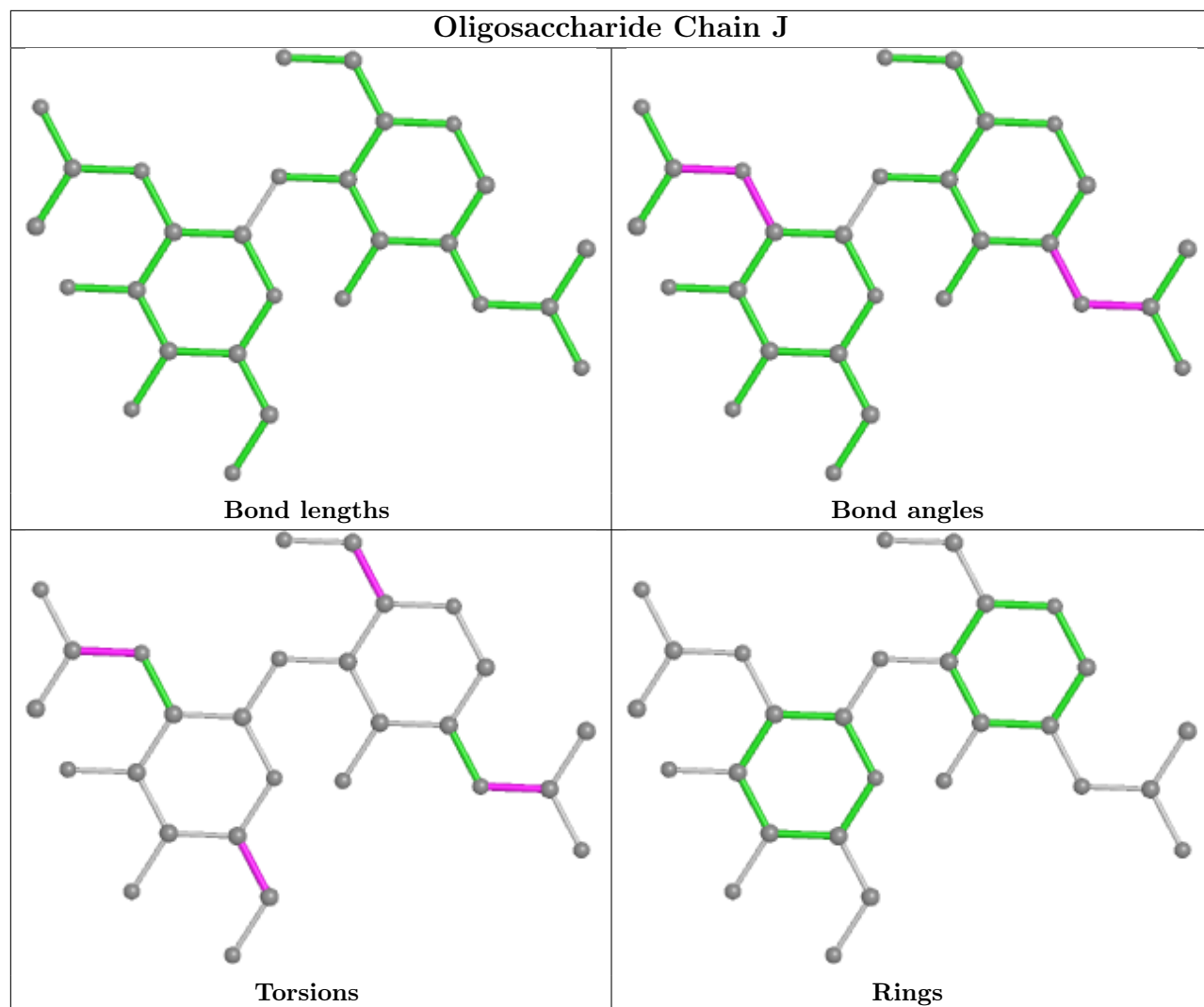
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	G	1	NAG	2	0
2	I	2	NAG	3	0
2	J	1	NAG	2	0
2	I	1	NAG	4	0
2	H	1	NAG	4	0
2	G	2	NAG	1	0
2	K	2	NAG	1	0
2	L	2	NAG	1	0
2	H	2	NAG	2	0
2	K	1	NAG	2	0
2	L	1	NAG	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

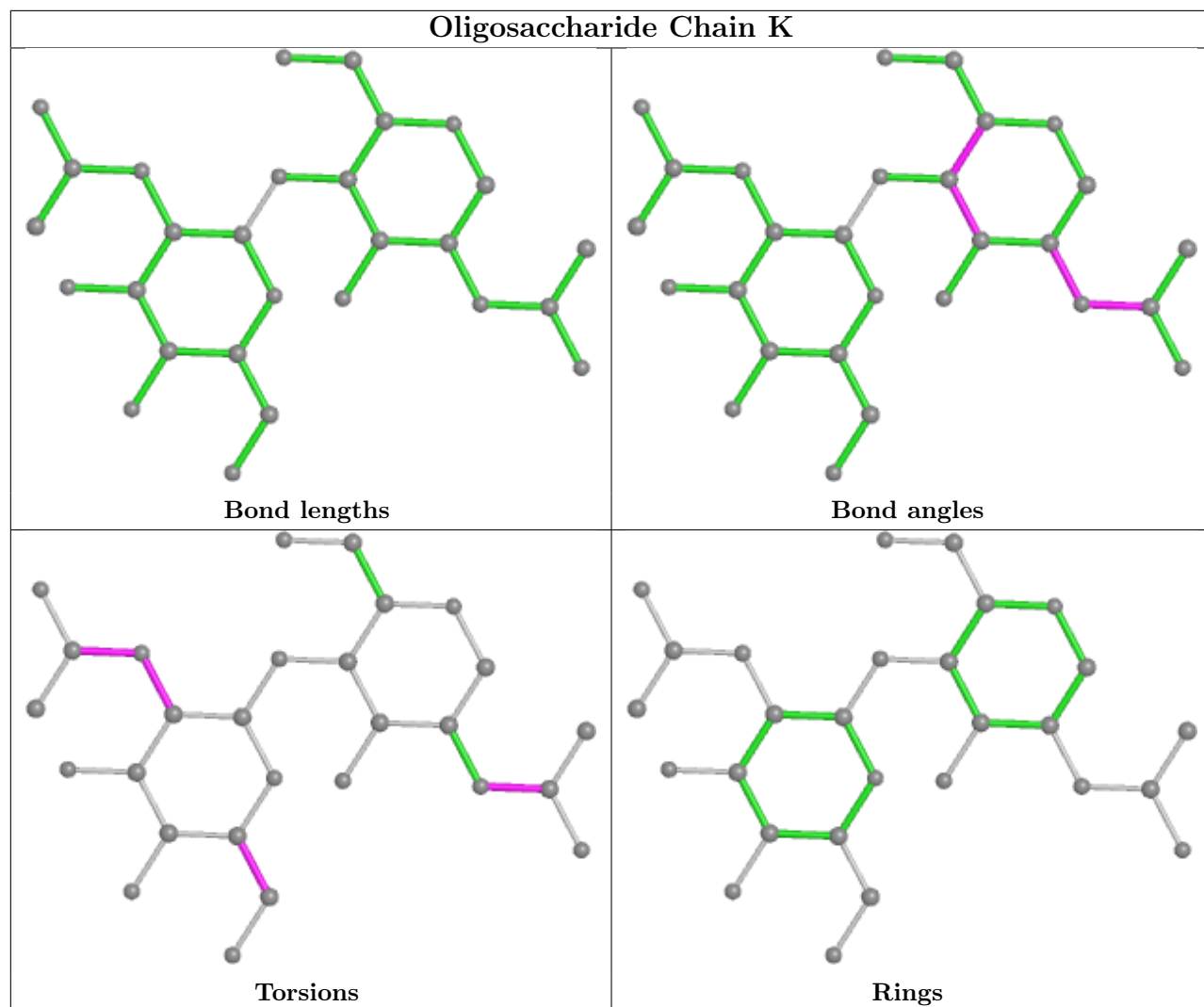


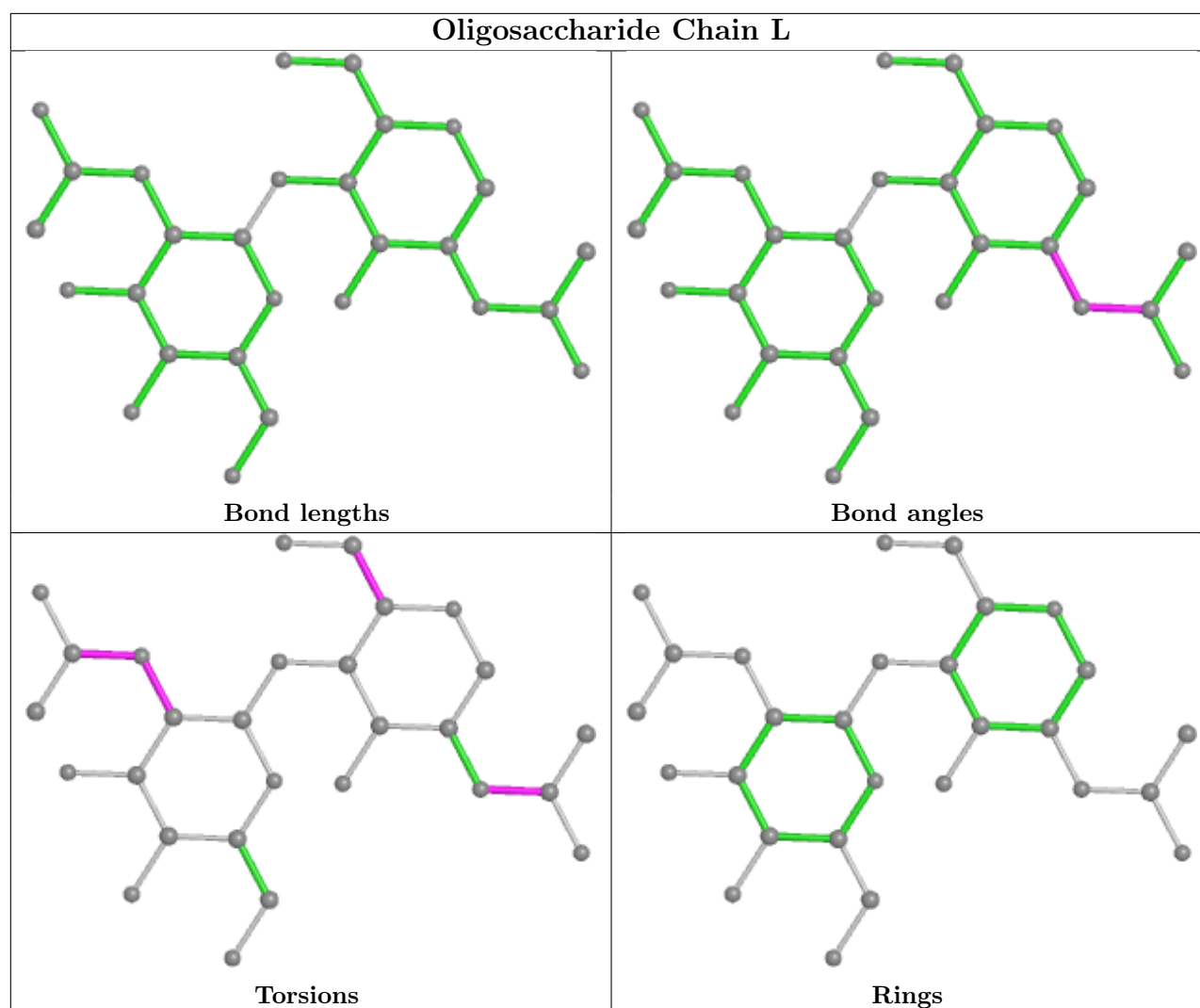












## 5.6 Ligand geometry [i](#)

27 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
4	NAG	B	702	1	14,14,15	0.62	0	17,19,21	0.65	0
5	EDO	C	801	-	3,3,3	0.59	0	2,2,2	0.43	0
5	EDO	A	802	-	3,3,3	0.49	0	2,2,2	0.40	0
5	EDO	F	802	-	3,3,3	0.50	0	2,2,2	0.40	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	NAG	A	702	1	14,14,15	0.69	0	17,19,21	0.67	0
4	NAG	F	703	1	14,14,15	0.59	0	17,19,21	0.60	0
3	GDL	B	600	-	15,15,15	3.97	3 (20%)	17,21,21	1.65	3 (17%)
3	GDL	D	600	-	15,15,15	3.81	5 (33%)	17,21,21	1.57	4 (23%)
5	EDO	A	801	-	3,3,3	0.60	0	2,2,2	0.43	0
5	EDO	A	800	-	3,3,3	0.57	0	2,2,2	0.52	0
5	EDO	F	801	-	3,3,3	0.59	0	2,2,2	0.45	0
3	GDL	A	600	-	15,15,15	3.69	4 (26%)	17,21,21	1.69	4 (23%)
5	EDO	B	800	-	3,3,3	0.58	0	2,2,2	0.47	0
5	EDO	E	802	-	3,3,3	0.69	0	2,2,2	0.40	0
5	EDO	E	801	-	3,3,3	0.58	0	2,2,2	0.38	0
3	GDL	E	600	-	15,15,15	3.96	3 (20%)	17,21,21	1.58	4 (23%)
5	EDO	C	800	-	3,3,3	0.56	0	2,2,2	0.46	0
3	GDL	C	600	-	15,15,15	3.76	4 (26%)	17,21,21	1.58	4 (23%)
4	NAG	F	702	1	14,14,15	0.60	0	17,19,21	0.76	1 (5%)
5	EDO	C	802	-	3,3,3	0.58	0	2,2,2	0.44	0
4	NAG	E	702	1	14,14,15	0.59	0	17,19,21	0.67	0
5	EDO	F	800	-	3,3,3	0.53	0	2,2,2	0.41	0
5	EDO	E	800	-	3,3,3	0.60	0	2,2,2	0.44	0
5	EDO	B	802	-	3,3,3	0.57	0	2,2,2	0.44	0
3	GDL	F	600	-	15,15,15	3.50	3 (20%)	17,21,21	1.66	4 (23%)
5	EDO	B	801	-	3,3,3	0.52	0	2,2,2	0.38	0
4	NAG	C	703	1	14,14,15	0.60	0	17,19,21	0.70	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	NAG	B	702	1	-	0/6/23/26	0/1/1/1
5	EDO	C	801	-	-	1/1/1/1	-
5	EDO	A	802	-	-	0/1/1/1	-
5	EDO	F	802	-	-	1/1/1/1	-
4	NAG	A	702	1	-	0/6/23/26	0/1/1/1
4	NAG	F	703	1	-	0/6/23/26	0/1/1/1
3	GDL	B	600	-	-	1/6/26/26	0/1/1/1
3	GDL	D	600	-	-	1/6/26/26	0/1/1/1
5	EDO	A	801	-	-	0/1/1/1	-
5	EDO	A	800	-	-	1/1/1/1	-
5	EDO	F	801	-	-	1/1/1/1	-

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	GDL	A	600	-	-	1/6/26/26	0/1/1/1
5	EDO	B	800	-	-	0/1/1/1	-
5	EDO	E	802	-	-	0/1/1/1	-
5	EDO	E	801	-	-	1/1/1/1	-
3	GDL	E	600	-	-	1/6/26/26	0/1/1/1
5	EDO	C	800	-	-	1/1/1/1	-
3	GDL	C	600	-	-	1/6/26/26	0/1/1/1
4	NAG	F	702	1	-	1/6/23/26	0/1/1/1
5	EDO	C	802	-	-	0/1/1/1	-
4	NAG	E	702	1	-	2/6/23/26	0/1/1/1
5	EDO	F	800	-	-	1/1/1/1	-
5	EDO	E	800	-	-	1/1/1/1	-
5	EDO	B	802	-	-	0/1/1/1	-
3	GDL	F	600	-	-	1/6/26/26	0/1/1/1
5	EDO	B	801	-	-	0/1/1/1	-
4	NAG	C	703	1	-	1/6/23/26	0/1/1/1

The worst 5 of 22 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	600	GDL	C2-C1	10.20	1.66	1.51
3	B	600	GDL	C3-C2	9.90	1.66	1.52
3	B	600	GDL	C2-C1	9.67	1.65	1.51
3	E	600	GDL	C3-C2	9.56	1.65	1.52
3	C	600	GDL	C2-C1	9.54	1.65	1.51

The worst 5 of 24 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	600	GDL	O5-C1-O1	3.96	124.26	118.47
3	B	600	GDL	O5-C1-O1	3.90	124.16	118.47
3	F	600	GDL	O5-C1-O1	3.83	124.06	118.47
3	C	600	GDL	O5-C1-O1	3.78	123.98	118.47
3	D	600	GDL	O5-C1-O1	3.78	123.98	118.47

There are no chirality outliers.

5 of 18 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	D	600	GDL	C1-C2-N2-C7
4	E	702	NAG	O5-C5-C6-O6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
4	E	702	NAG	C4-C5-C6-O6
5	A	800	EDO	O1-C1-C2-O2
5	C	801	EDO	O1-C1-C2-O2

There are no ring outliers.

7 monomers are involved in 13 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	C	801	EDO	2	0
4	A	702	NAG	1	0
5	A	801	EDO	2	0
5	A	800	EDO	2	0
5	E	801	EDO	1	0
5	F	800	EDO	1	0
5	B	801	EDO	4	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	484/515 (93%)	-0.04	15 (3%) 49 52	27, 38, 58, 85	0
1	B	484/515 (93%)	-0.15	10 (2%) 63 66	24, 37, 57, 85	0
1	C	484/515 (93%)	-0.05	19 (3%) 39 42	27, 38, 67, 88	0
1	D	483/515 (93%)	-0.03	14 (2%) 51 55	27, 39, 64, 93	0
1	E	484/515 (93%)	-0.11	16 (3%) 46 48	25, 36, 54, 86	0
1	F	484/515 (93%)	-0.05	24 (4%) 28 31	28, 39, 66, 90	0
All	All	2903/3090 (93%)	-0.07	98 (3%) 45 47	24, 38, 62, 93	0

The worst 5 of 98 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	554	GLU	13.6
1	C	554	GLU	10.2
1	B	554	GLU	9.1
1	D	317	ASP	7.8
1	E	315	LYS	7.5

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

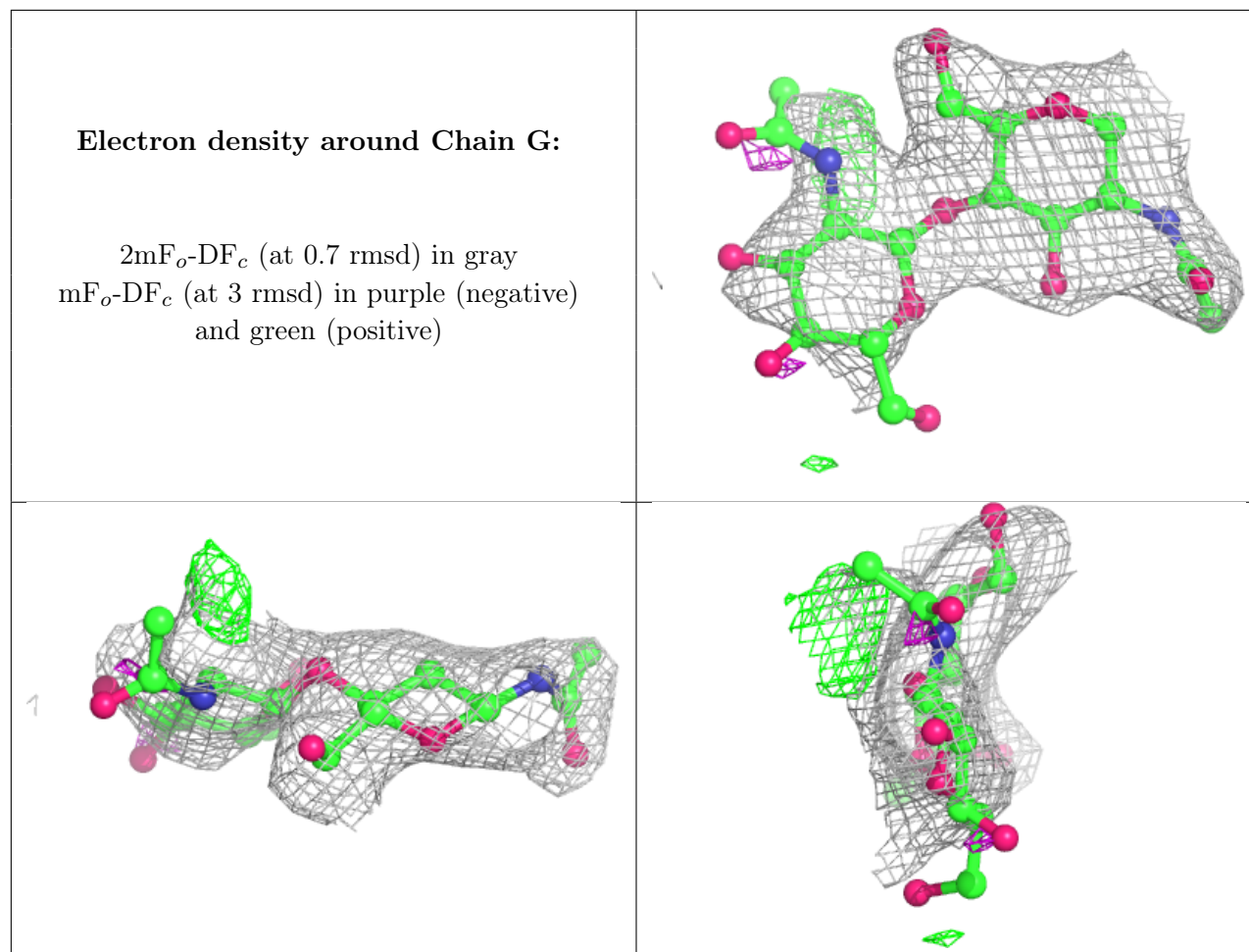
There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	NAG	H	2	14/15	0.60	0.47	80,84,87,90	0
2	NAG	G	2	14/15	0.71	0.56	89,93,95,96	0
2	NAG	I	2	14/15	0.72	0.56	85,88,90,91	0
2	NAG	J	2	14/15	0.76	0.42	81,83,86,88	0
2	NAG	K	2	14/15	0.78	0.46	81,84,86,87	0
2	NAG	L	2	14/15	0.82	0.54	80,85,88,90	0
2	NAG	L	1	14/15	0.84	0.25	57,60,66,74	0
2	NAG	K	1	14/15	0.85	0.26	64,67,70,76	0
2	NAG	J	1	14/15	0.85	0.21	64,68,72,77	0
2	NAG	H	1	14/15	0.86	0.18	62,65,69,75	0
2	NAG	G	1	14/15	0.88	0.17	66,70,74,82	0
2	NAG	I	1	14/15	0.89	0.24	66,69,74,80	0

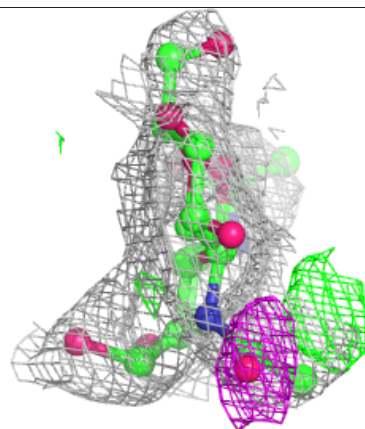
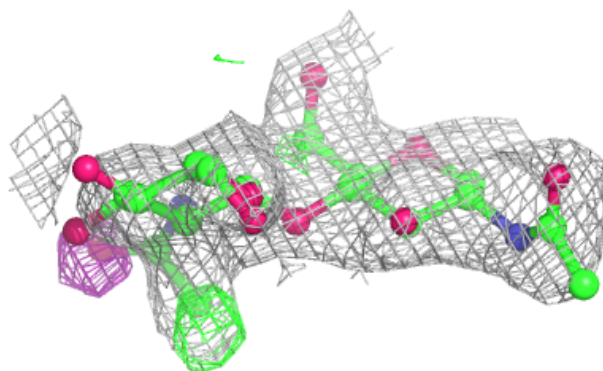
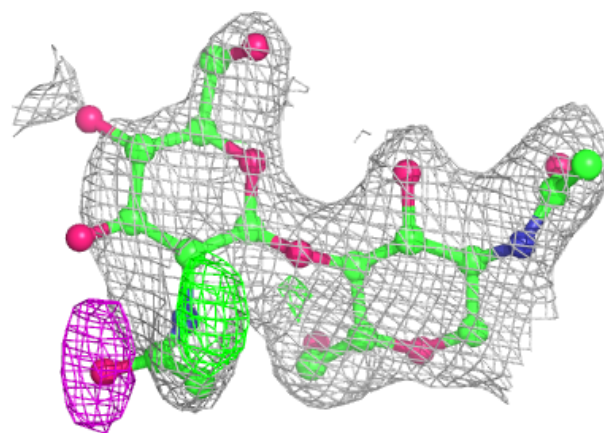
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.



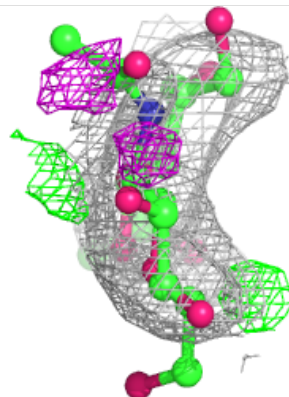
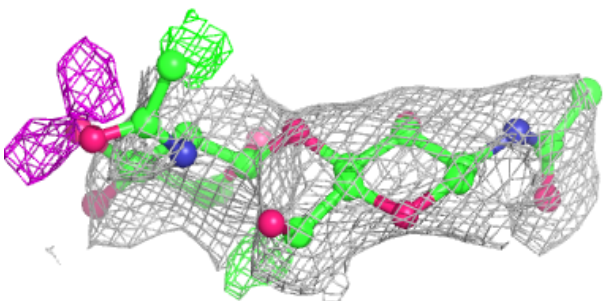
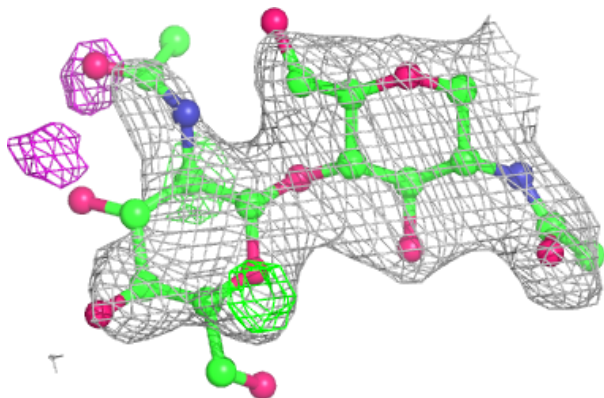


**Electron density around Chain H:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

**Electron density around Chain I:**

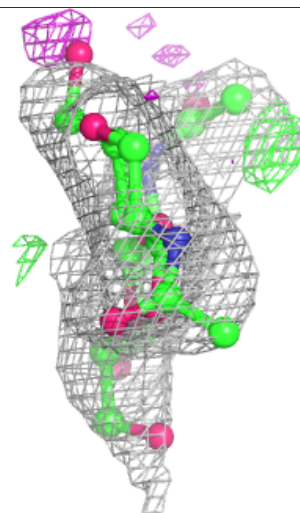
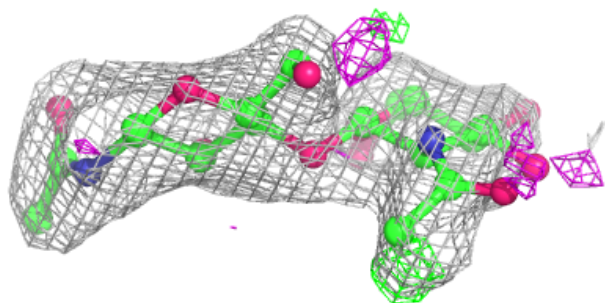
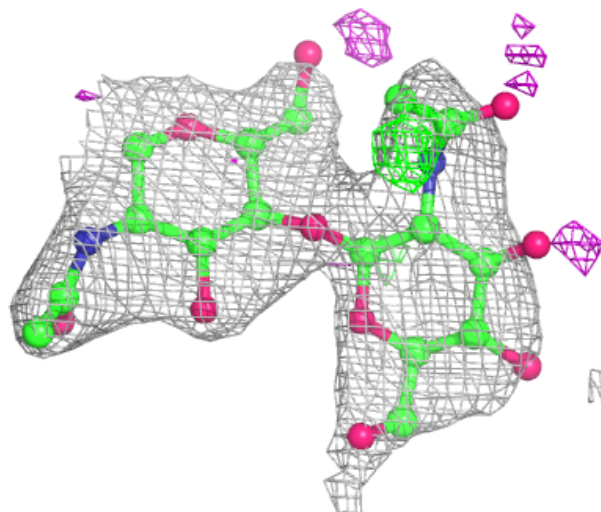
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





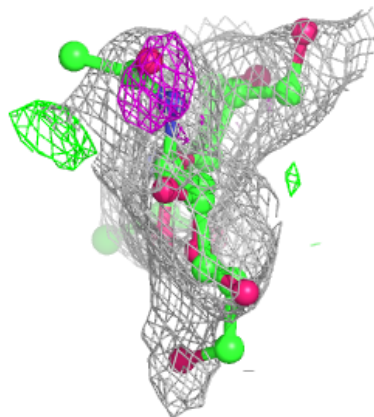
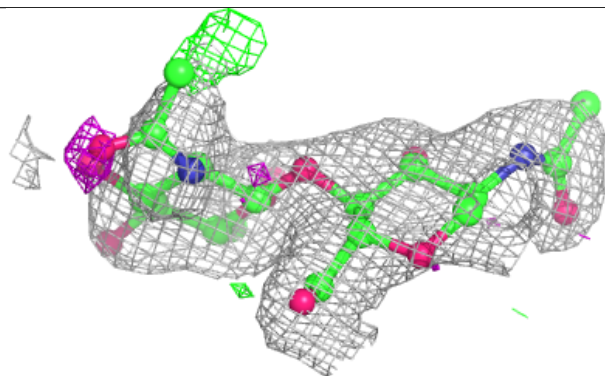
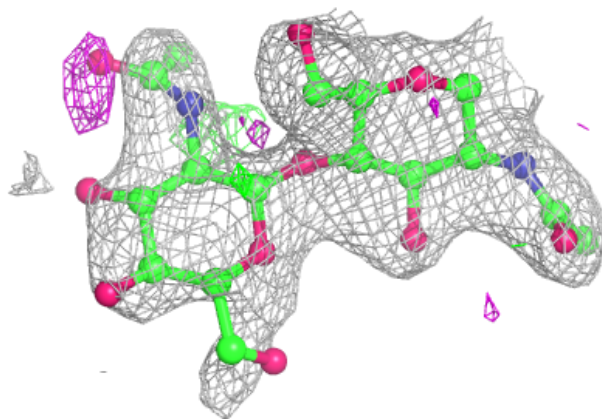
**Electron density around Chain J:**

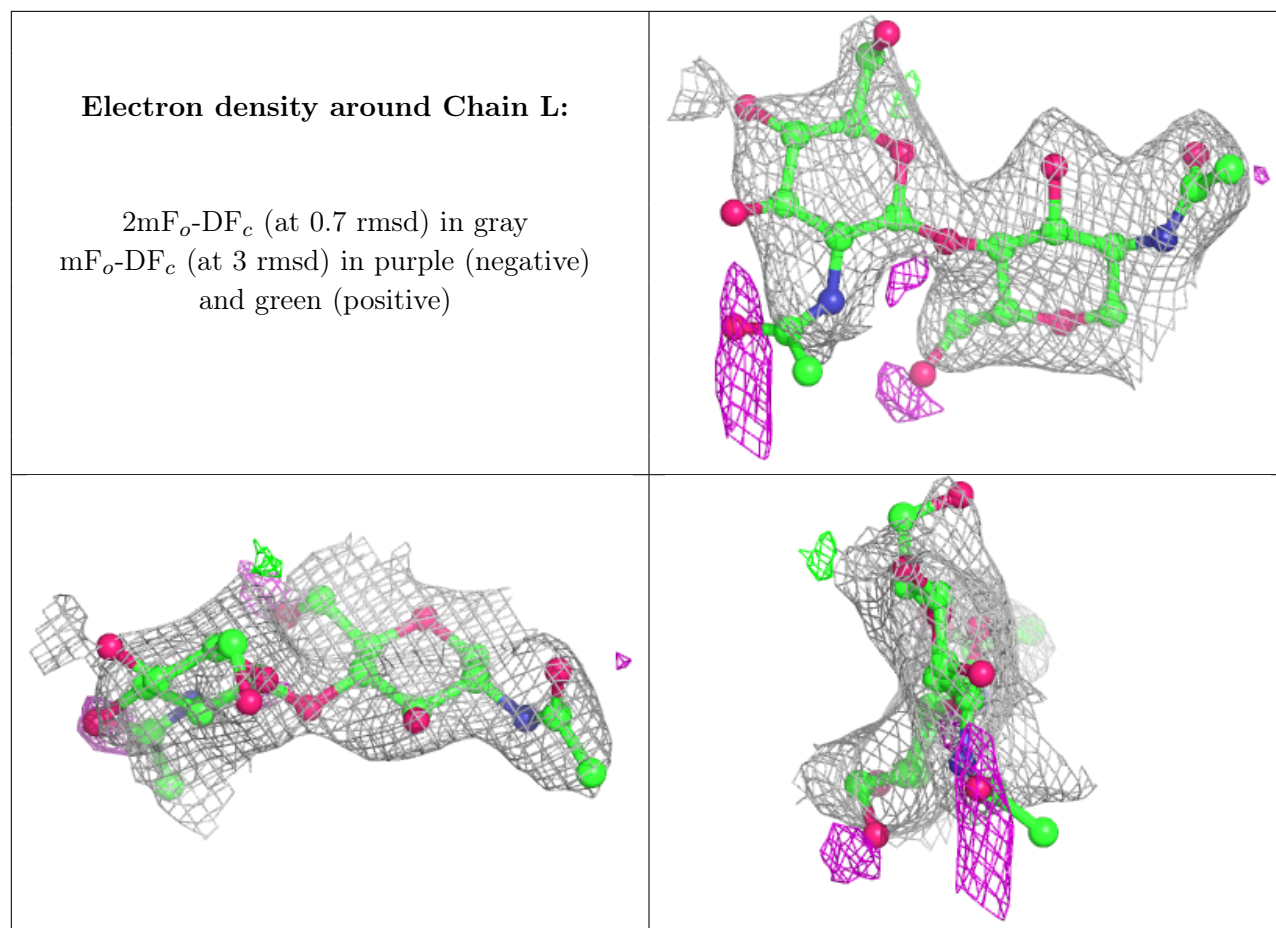
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around Chain K:**

$2mF_o - DF_c$  (at 0.7 rmsd) in gray  
 $mF_o - DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	NAG	B	702	14/15	0.67	0.40	70,72,75,76	0
4	NAG	E	702	14/15	0.68	0.48	77,81,82,83	0
4	NAG	C	703	14/15	0.69	0.54	84,88,90,90	0
4	NAG	A	702	14/15	0.70	0.34	64,71,75,78	0
4	NAG	F	702	14/15	0.70	0.52	83,88,90,91	0
5	EDO	C	801	4/4	0.77	0.30	65,66,67,67	0
4	NAG	F	703	14/15	0.80	0.45	79,82,84,84	0
5	EDO	B	801	4/4	0.84	0.31	61,62,63,65	0
5	EDO	A	801	4/4	0.84	0.35	76,76,76,77	0
5	EDO	E	801	4/4	0.89	0.27	65,66,66,67	0
3	GDL	F	600	15/15	0.90	0.10	33,38,40,41	0
5	EDO	C	800	4/4	0.91	0.19	59,59,59,61	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	EDO	E	802	4/4	0.91	0.16	30,31,34,38	0
5	EDO	F	800	4/4	0.91	0.17	48,51,52,56	0
5	EDO	F	801	4/4	0.91	0.31	66,67,67,69	0
5	EDO	E	800	4/4	0.92	0.13	49,49,49,51	0
3	GDL	C	600	15/15	0.93	0.10	30,34,38,38	0
5	EDO	B	802	4/4	0.93	0.17	30,35,36,38	0
3	GDL	E	600	15/15	0.93	0.12	29,33,34,36	0
3	GDL	A	600	15/15	0.94	0.10	30,32,34,35	0
3	GDL	D	600	15/15	0.94	0.11	32,34,37,40	0
3	GDL	B	600	15/15	0.95	0.10	30,33,34,35	0
5	EDO	F	802	4/4	0.95	0.17	37,38,39,39	0
5	EDO	A	802	4/4	0.96	0.17	31,38,40,40	0
5	EDO	A	800	4/4	0.97	0.11	56,57,57,61	0
5	EDO	B	800	4/4	0.97	0.12	38,38,38,41	0
5	EDO	C	802	4/4	0.98	0.08	35,36,37,38	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.